TOPICS IN LOGIC AND APPLICATIONS

ANUSH TSERUNYAN

These lecture notes encompass the author's three-day course given at the 2016 Undergraduate Summer School on Model Theory at University of Notre Dame.

Section 1 contains an introduction to ideals and filters, as well as finitely additive measures in general; it also includes ultrafilters and a proof of Hindman's theorem.

In Section 2, we present the construction of ultraproducts, together with Łoś's theorem, and give a proof of the Compactness theorem via ultraproducts. We also discuss several combinatorial and measure-theoretic applications of the Compactness theorem. The last subsection introduces the concept of saturation and ends with a proof of it for ultraproducts.

In Section 3, we investigate the structure of a nonstandard extension $\mathbb{R}^*$ of $\mathbb{R}$ and give several nonstandard characterizations of basic concepts from real analysis, as well as nonstandard proofs of familiar theorems of calculus.

The notes end with a list of exercises that go along with the material.

Contents

Acknowledgments. The author thanks Gabriel Conant, Anand Pillay, and Sergei Starchenko, for all their input and help. She is also thankful to Mike Haskel and Gregory Cousins for their comments and corrections.

Credits. Section 2 of these notes uses material from [Tse13] and Section 3 closely follows [Gol], in fact, some parts are verbatim.

1. Ultrafilters

1.A. Ideals and filters

Let's recall Cantor's proof of the existence of transcendental numbers. The idea was that $\mathbb{Q}$ is a very small subset of $\mathbb{R}$; it is, in fact, so small that even the set of all roots of all polynomials over $\mathbb{Q}$ is still as small as $\mathbb{Q}$, and hence its complement (the set of transcendental numbers) is large and, in particular, nonempty! The *notion of smallness* that makes all these statements true is being *countable*.

This has become a powerful method for proving the existence of certain kinds of elements, temporarily call them "good", in a given nonempty set X (in the above example, $X := \mathbb{R}$): one introduces a notion of smallness (or equivalently, largeness) of subsets of X and shows that the set of "good" elements is large. The proofs of the latter statement often require our notion of smallness to be sufficiently additive, i.e. the union of two small sets is still small; in fact, it is crucial in Cantor's proof that countable union of small sets is still small. The following definition isolates such notions of smallness.

Definition 1.1. An *ideal* on a set X is a nonempty collection $\mathcal{I} \subseteq \mathscr{P}(X)$ that is

(i) closed *downward*: $B \subseteq A \in \mathcal{I} \implies B \in \mathcal{I}$,

(ii) closed under *finite unions*: $A, B \in \mathcal{I} \implies A \cup B \in \mathcal{I}$,

(iii) nontrivial: $X \notin \mathcal{I}$.

An ideal $\mathcal{I}$ is called a σ-*ideal* if (ii) is strengthened to

(ii′) closed under *countable unions*: $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{I} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{I}$.

Note that every ideal contains $\varnothing$. Also, by induction, condition (ii) is equivalent to $\{A_n\}_{n<N} \subseteq \mathcal{I} \implies \bigcup_{n<N} A_n \in \mathcal{I}$, for any $N \in \mathbb{N}$.

Examples 1.2.

(a) The collection $\mathcal{I}_F$ of finite subset of an infinite set X is an ideal, called the *Fréchet ideal* on X.

(b) The collection of countable subsets of an uncountable set is a σ-ideal.

(c) The collection of nowhere dense[1] subsets of a nonempty topological space (e.g. a metric space) is an ideal.

(d) A set $A \subseteq \mathbb{N}$ is called *summable* if $\sum_{n\in A} \frac{1}{n} < \infty$. It is clear that the collection of summable sets forms an ideal, called the *summable ideal*.

Taking the complements of sets in an ideal, we get a dual notion of largeness, explicitly stated in the following definition.

Definition 1.3. An *filter* on a set X is a nonempty collection $\mathcal{F} \subseteq \mathscr{P}(X)$ that is

(i) closed *upward*: $B \supseteq A \in \mathcal{F} \implies B \in \mathcal{F}$,

(ii) closed under *finite intersections*: $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$,

(iii) *nontrivial*: $\varnothing \notin \mathcal{F}$.

[1] A subset A of a topological space X is called *nowhere dense* if every nonempty open set U has a further nonempty open $V \subseteq U$ disjoint from A. This is equivalent to the closure $\overline{A}$ not containing a nonempty open set.

A filter $\mathcal{F}$ is called a *δ-filter* if (ii) is strengthened to

(ii′) closed under *countable intersections*: $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F} \implies \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

For a collection $\mathcal{C} \subseteq \mathscr{P}(X)$, put $\mathcal{C}' := \{A^c : A \in \mathcal{C}\}$ and call it the *dual* of $\mathcal{C}$.

Proposition 1.4. *The dual of an ideal is a filter and vice versa.*

Proof. Straightforward verification. □

Thus, the aforementioned examples of ideals define corresponding filters: the *Fréchet filter* of cofinite sets, the δ-filter of cocountable sets, the summable filter, and the filter of co-nowhere-dense sets.

Example 1.5. Take any point $x \in X$ and give that point the full mass, i.e. define a filter δ_x by putting a set $A \subseteq X$ in δ_x if and only if $A \ni x$. This is indeed a filter, but it's not useful at all because it only "sees" the one point x and reduces all of the statements about subsets of X to those about x. Filters of the form δ_x are called *principal*, so a filter is *nonprincipal* if it does not contain any singleton, and hence, any finite set.

Lastly, we give an important example of a filter that is vastly used in arithmetic combinatorics and ergodic theory.

Example 1.6. Define a partial function $d : \mathscr{P}(\mathbb{N}) \rightharpoonup [0,1]$ by

$$d(A) := \lim_{n\to\infty} \frac{|A \cap [0,n)|}{n},$$

whenever the limit exists. Call $d(A)$ the *density* of A. The sets for which the density is defined and is equal to 1 form a filter called the *density filter*.

Terminology and notation. $\mathcal{F}$ be a filter on a set X. We call a set $A \subseteq X$

- *$\mathcal{F}$-large* if $A \in \mathcal{F}$,
- *$\mathcal{F}$-small* if $A \in \mathcal{F}'$,
- *$\mathcal{F}$-intermediate* if A is neither $\mathcal{F}$-large nor $\mathcal{F}$-small,
- *$\mathcal{F}$-positive* if A is not $\mathcal{F}$-small.

Caution 1.7. "Not small" does not mean "large". Indeed, one can easily exhibit intermediate sets for each of the aforementioned examples of filter/ideals.

For a property P of elements of X, we say that P *holds $\mathcal{F}$-almost-everywhere* (write $\mathcal{F}$-a.e.) or *for $\mathcal{F}$-a.e.* $x \in X$ if the set $\{x \in X : P(x)\}$ is $\mathcal{F}$-large. Symbolically, this is written

$$\forall^{\mathcal{F}} x \in X \; P(x).$$

We also write

$$\exists^{\mathcal{F}} x \in X \; P(x)$$

to mean that the set $\{x \in X : P(x)\}$ is $\mathcal{F}$-positive. Note that the analogues of De Morgan's laws still hold, e.g., $\neg\forall^{\mathcal{F}} = \exists^{\mathcal{F}}\neg$.

1.B. Finitely additive measures

The notions of ideal and filter can be unified into that of *measure*, to define which we first need the following.

Definition 1.8. An *algebra* on a set X is a nonempty collection $\mathcal{A} \subseteq \mathscr{P}(X)$ that is

(i) closed under *complements*: $A \in \mathcal{A} \implies A^c \in \mathcal{A}$,

(ii) closed under *finite unions*: $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$.

A algebra $\mathcal{A}$ is called a *σ-algebra* if (ii) is strengthened to

(ii′) closed under *countable unions*: $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Proposition 1.9. *Every algebra (resp. σ-algebra) contains ∅ and X, and is closed under finite (resp. countable) intersections.*

Proof. Easy, left as an exercise. □

Proposition 1.10. *If $\mathcal{I}$ is an ideal then $\mathcal{I} \cup \mathcal{I}'$ is an algebra.*

Proof. Straightforward verification. □

Notation 1.11. We use the symbol $\bigsqcup$ to denote the union of pairwise disjoint sets. Thus, $A = \bigsqcup_{i\in I} A_i$ means that A is equal to $\bigcup_{i\in I} A_i$ and the sets A_i are pairwise disjoint.

Example 1.12. Let $\mathcal{A}$ be a collection of disjoint finite unions of intervals in $\mathbb{R}$, i.e. each element of $\mathcal{A}$ is of the form $\bigsqcup_{n<k} I_n$, where each I_n is an interval[2]. It is not hard to check that $\mathcal{A}$ forms an algebra on $\mathbb{R}$.

Definition 1.13. A *finitely additive* (f.a.) *measure* on an algebra $\mathcal{A} \subseteq \mathscr{P}(X)$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ that is

(i) *finitely additive*: if $A, B \in \mathcal{A}$ are disjoint, then $\mu(A \sqcup B) = \mu(A) + \mu(B)$,

(ii) $\mu(\varnothing) = 0$.

An f.a. measure μ is called just a *measure* (or a *countably additive measure*) if (i) is strengthened to

(i′) *countably additive*: if $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ is pairwise disjoint, then $\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

For an f.a. measure μ on an algebra $\mathcal{A}$, a set $A \subseteq X$ is called *μ-null* if it is a subset of a set $B \in \mathcal{A}$ with $\mu(B) = 0$. Consequently, a set $A \subseteq X$ is called *μ-conull* if its complement is μ-null. We say that μ is *complete* if $\mathcal{A}$ contains all of the μ-null sets. Requiring an f.a. measure to be complete is not restrictive because any f.a. measure can be easily completed by extending it to the algebra generated by $\mathcal{A}$ and the μ-null sets.

Proposition 1.14. *The μ-null sets of an f.a. measure form an ideal. Conversely, for any ideal $\mathcal{I}$, there is a unique complete f.a. measure $\mu_{\mathcal{I}} : \mathcal{I} \cup \mathcal{I}' \to \{0, 1\}$ whose null sets are exactly those in $\mathcal{I}$.*

Proof. Define $\mu(A)$ to be 0 if $A \in \mathcal{I}$, and 1 if $A \in \mathcal{I}'$. □

[2] Here, by *interval* we mean sets of the form $(a, b) \cup C$, where $a \leq b$ and $C \subseteq \{a, b\}$.

Thus, one can think of ideals and filters as the collections of μ-null and μ-conull sets of $\{0,1\}$-valued complete f.a. measures.

Examples 1.15.

(a) For $\mathcal{A}$ be as in Example 1.12, we define an f.a. measure ν on $\mathcal{A}$ by defining, for each $\bigsqcup_{n<k} I_n \in \mathcal{A}$,

$$\nu\left(\bigsqcup_{n<k} I_n\right) := \sum_{n<k} |I_n|,$$

where I_n denotes the length of the interval I_n. It is easy to check that ν is indeed an f.a. measure.

(b) Extending the previous example, the Lebesgue measure λ on $\mathbb{R}$, or more generally on $\mathbb{R}^n$, is a (countably additive) complete measure.

1.C. **Applications**

Ideals and filters are used to prove existence of not only individual objects, but also arbitrarily large or even infinite sets of objects. A toy example is the Infinite Pigeonhole Principle (IPHP), which states the following:

Infinite Pigeonhole Principle 1.16. *If an infinite set is partitioned into finitely many sets, then one of those sets must be infinite.*

The obvious proof of this fact is equivalent to the statement that the collection of finite subsets of an infinite set forms an ideal, namely, the Fréchet ideal. This principle is true for any ideal in general:

Pigeonhole Principle for ideals 1.17. *Let $\mathcal{I}$ be an ideal on a set X. If an $\mathcal{I}$-positive set A is partitioned into finitely many sets, then one of those sets is again $\mathcal{I}$-positive.*

Using the IPHP, one can immediately prove the following:

König's Lemma 1.18. *Any infinite locally finite connected graph contains an infinite simple path*[3]*.*

Proof. By taking a spanning subtree, we may assume without loss of generality that our graph is a tree to begin with. Fix a vertex v_0, call it a root of the tree and direct all of the edges away from v_0, i.e. for each vertex v, the unique path connecting v_0 and v is a directed ath from v_0 to v. If (u,v) is a directed edge, call v a child of u.

For a vertex v, denote by $A(v)$ the set of all *ancestors* of v, i.e. all vertices such that the unique directed path connecting them to v starts from v. We make a convention that $A(v)$ also contains v itself.

By recursion, we build a simple path $(v_n)_{n\in\mathbb{N}}$ such that $A(v_n)$ is infinite for each $n \in \mathbb{N}$. Starting from v_0, assume that we have already built a desired path $(v_n)_{n\leq k}$. Since our graph is locally finite, v_k has only finitely many children $u_1, u_2, ..., u_m$. Because $A(v_k) = \{v_k\} \cup \bigcup_{i\leq m} A(u_i)$ and $A(v_k)$ is infinite by the inductive hypothesis, one of $A(u_i)$ is infinite. Choose one such u_i and that will be our v_{k+1}. □

[3]An path in a graph is called *simple* if all vertices along it are pairwise distinct (no vertex appears more than once)

IPHP can be amplified to an extremely useful 2-dimensional version known as the Infinite Ramsey theorem.

For a set S, let $[S]^2$ denote the set of two element subsets of S (think of it as the set of edges of the undirected complete graph on S). Given a finite coloring χ of $[\mathbb{N}]^2$, i.e. a function $\chi : [\mathbb{N}]^2 \to \{0, 1, \cdots, k\}$ for some $k \in \mathbb{N}$, an edge-set $E \subseteq [\mathbb{N}]^2$ is said to be *χ-monochromatic* if all elements of E have the same color, i.e. $\chi\downharpoonright_E$ is constant. A vertex-set $A \subseteq \mathbb{N}$ is called *χ-monochromatic* if $[A]^2$ is monochromatic.

Infinite Ramsey Theorem 1.19. *For any finite coloring χ of $[\mathbb{N}]^2$, there exists an infinite χ-monochromatic subset of $\mathbb{N}$.*

Proof. The idea is we use the IPHP to produce a finite coloring c of vertices out of the given finite coloring χ of the edges; then we apply the IPHP again to obtain a c-monochromatic set, so the IPHP gets used twice (which is a reasonable cost to pay for switching from 2 dimensions to 1).

For $a \in \mathbb{N}$ and $A \subseteq \mathbb{N}$, put $(a, A) := \{\{a, a'\} : a' \in A \setminus \{a\}\}$. Set $A_0 := \mathbb{N}$ and take sequences $a_n \in \mathbb{N}$ and $A_n \subseteq \mathbb{N}$ satisfying:

(i) $a_n \in A_n$,

(ii) $A_{n+1} \subseteq A_n$ is infinite and (a_n, A_{n+1}) is χ-monochromatic.

It is easy to see that such sequences $(a_n)_{n\in\mathbb{N}}$ and $(A_n)_{n\in\mathbb{N}}$ exist (define them recursively using the IPHP). Define a finite coloring c on $\{a_n\}_{n\in\mathbb{N}}$ by coloring a_n with the common χ-color of all edges in (a_n, A_{n+1}). By the IPHP again, there is a c-monochromatic infinite subsequence $(a_{n_k})_{k\in\mathbb{N}}$. Now it is straightforward to check that $A := \{a_{n_k}\}_{k\in\mathbb{N}}$ is χ-monochromatic. □

1.D. **Ultrafilters and applications**

The use of filters in proofs can often be hard or not work at all due to the existence of intermediate sets. Having a filter at hand for which “not small” means “ not only makes many arguments shorter and more conceptual/elegant, but also enables new tools yielding strong and surprising theorems.

Definition 1.20. A filter $\mathcal{F}$ on a set X is called an *ultrafilter* if $\mathcal{F} \cup \mathcal{F}' = \mathscr{P}(X)$.

We use lowercase Greek letters $\alpha, \beta, \gamma, \cdots$ to denote ultrafilters. Note that the defining property of ultrafilters α is that

$$\exists^\alpha = \forall^\alpha,$$

making De Morgan law’s look strange:

$$\neg\forall^\alpha = \forall^\alpha\neg.$$

Examples of ultrafilters? Well, any principal filter is an ultrafilter, but, as mentioned above, these are not interesting examples. However, it is not even clear that there are nonprincipal ultrafilters. Turns out their existence follows from Axiom of Choice and is independent from ZF, so they cannot be defined constructively.

Lemma 1.21 (Uses Axiom of Choice)**.** *Every filter is contained in an ultrafilter. In particular, if X is an infinite set, the Fréchet filter is contained in an ultrafilter, which hence is nonprincipal.*

Proof. Zorn’s lemma. □

To illustrate the use of ultrafilters, we will now prove the following well-known (and super-cool) theorem using a special kind of ultrafilters on $\mathbb{N}$. To state it we need the following notion.

Definition 1.22. For a set $A \subseteq \mathbb{N}$, the *finite-sums set generated by* A is the set

$$\Sigma(A) := \left\{ \sum_{n \in F} n : F \text{ is a finite subset of } A \right\}.$$

A set $P \subseteq \mathbb{N}$ is called an *IP set* if it is the finite-sums set generated by an infinite $A \subseteq \mathbb{N}$.

Note that for a $A \subseteq \mathbb{N}$, $\Sigma(A)$ is finite if and only if A is finite, so IP sets are infinite by definition.

Remark 1.23. The term IP stands for *infinite-dimensional parallelepiped*. It is due to the illustration of $\Sigma(A)$ as the set of nonzero vertices of the parallelepiped based at 0 with edges being the elements of A viewed as orthogonal vectors originating at 0.

Theorem 1.24 (Hindman)**.** *Whenever* $\mathbb{N}$ *is partitioned into finitely many sets, one of these sets contains an infinite IP set.*

To get ready for the proof we need some notation and an important definition.

For an ultrafilter α and a set $A \subseteq \mathbb{N}$, put

$$\Delta_\alpha(A) := \{d \in \mathbb{N} : A - d \text{ is } \alpha\text{-large}\},$$

where

$$A - d := \{n \in \mathbb{N} : n + d \in A\},$$

in other words, $A - d$ is the inverse image of the function $n \mapsto n + d$.

Definition 1.25. An ultrafilter α on $\mathbb{N}$ is called *idempotent* if for any $A \subseteq \mathbb{N}$,

$$A \text{ is } \alpha\text{-large} \iff \Delta_\alpha(A) \text{ is } \alpha\text{-large};$$

symbolically, A is α-large $\iff$ $(\forall^\alpha d \in \mathbb{N})\ A - d$ is α-large.

Clearly, idempotent ultrafilters on $\mathbb{N}$ have to be nonprincipal and it's not at all clear that such ultrafiltes exist. However, the following is a corollary of a more general theorem of Ellis. We'll skip the proof of it, which is based on a clever application of Zorn's lemma.

Proposition 1.26 (Axiom of Choice)**.** *There are idempotent ultrafilters on* $\mathbb{N}$.

We need one more piece of notation: for $A \subseteq \mathbb{N}$ and $k \in \mathbb{N}$, put

$$\partial_k A := A \cap (A - k);$$

one can think of $\partial_k A$ as the *directional derivative of* A *in the direction* k, hence the notation.

Proof of Theorem 1.24. Fix an idempotent ultrafilter α on $\mathbb{N}$. Given a finite partition of $\mathbb{N}$, exactly one of the sets in the partition is α-large; denote it by A_0. By recursion, we define a sequence $(A_n)_{n \in \mathbb{N}}$ of α-large decreasing sets together with a sequence $(k_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ such that $\partial_{k_n} A_n \supseteq A_{n+1}$, so $k_n + A_{n+1} = A_n$.

Assume the set A_n and the sequence $(k_i)_{i<n}$ is defined. By idempotence, $\Delta_\alpha(A_n)$ is α-large, so $A'_n := A_n \cap \Delta_\alpha(A_n)$ is also α-large. Because α is nonprincipal, the set $A'_n \setminus \{k_i\}_{i<n}$ is still α-large and hence nonempty, so we let k_n be an arbitrary element of the latter set. Finally, put $A_{n+1} := \partial_{k_n} A_n$.

Put $A \coloneqq \{k_n\}_{n\in\mathbb{N}}$ and note that A is infinite by the choice of the k_n, so $\Sigma(A)$ is an IP set. Furthermore, one can easily verify by induction on $l \in \mathbb{N}$ that for any $n_1 < n_2 < \cdots < n_l$, $k_{n_1} + k_{n_2} + \cdots + k_{n_l} \in A_{n_1}$. Thus, $\Sigma(A) \subseteq A_0$. $\square$

2. Ultraproducts and compactness

2.A. The construction and Łoś's theorem

Let I be an index set (possibly uncountable) and let α be a nonprincipal ultrafilter on I. For a sequence $(X_i)_{i\in I}$ of sets, we think of elements x, y of the product $\prod_{i\in I} X_i$ as functions $x, y : I \to \bigcup_{i\in I X_i}$, and thus, define the following equivalence relation

$$x =_\alpha y :\iff x(i) = y(i) \text{ for } \alpha\text{-a.e. } i \in I$$

just like we do with functions on measure spaces (e.g. the L^p spaces). We call the quotient space $X_\infty \coloneqq \prod_{i\in I} X_i / =_\alpha$ the *ultraproduct of* $(X_i)_{i\in I}$ *along the ultrafilter* α; we will use the notation $\prod_{i\in I} X_i/\alpha$ instead, omitting $=$. Continuing the analogy with the usual L^p spaces, we identify $x \in \prod_{i\in I} X_i$ with its equivalence class $[x]_\alpha$; likewise, we often identify a subset S of $\prod_{i\in I} X_i$ with the union $[S]_\alpha$ of the equivalence classes of the elements of S.

Notation 2.1 (for vectors). For a vector $\vec{x} = (x^{(1)}, x^{(2)}, \cdots, x^{(n)}) \in \left(\prod_{i\in I} X_i\right)^n$, put

$$\vec{x}(i) \coloneqq (x^{(1)}(i), x^{(2)}(i), \cdots, x^{(n)}(i)),$$

for $i \in I$, and

$$[\vec{x}]_\alpha \coloneqq ([x^{(1)}]_\alpha, [x^{(2)}]_\alpha, \cdots, [x^{(n)}]_\alpha).$$

One can think of the ultraproduct as a limit of the sets X_i, and, as such, it inherits the properties and structure enjoyed by α-a.e. X_i. For example, if each X_i is actually a group $(G_i, e_i, \cdot_i)$, then we can turn their ultraproduct into a group: simply define the multiplication coordinate-wise and $(e_i)_{i\in I}$ would be the identity. This is true more generally.

Definition 2.2 (Ultraproduct of structures)**.** Let $\mathcal{L}$ be a first-order language and $(\mathcal{M}_i)_{i\in I}$ a (possibly uncountable) sequence of $\mathcal{L}$-structures. We define the *ultraproduct* $\mathcal{L}$*-structure* $\mathcal{M}_\infty$ *of* $(\mathcal{M}_i)_{i\in I}$ *along an ultrafilter* α as follows:

(i) Universe: let M_∞ be the ultraproduct of sets $(M_i)_{i\in I}$;

(ii) Constants: for each constant symbol c in $\mathcal{L}$, put
$$c^{\mathcal{M}_\infty} \coloneqq [(c^{\mathcal{M}_i})_{i\in I}];$$

(iii) Functions: for each function symbol f in $\mathcal{L}$ with arity n and for each vector $\vec{a} \in \left(\prod_{i\in I} M_i\right)^n$,
$$f^{\mathcal{M}_\infty}([\vec{a}]_\alpha) \coloneqq [f^{\mathcal{M}_i}(\vec{a}(i))]_\alpha.$$

(iv) Relations: for each relation symbol R in $\mathcal{L}$ with arity n and for each vector $\vec{a} \in \left(\prod_{i\in I} M_i\right)^n$,
$$R^{\mathcal{M}_\infty}([\vec{a}]_\alpha) :\Leftrightarrow (\forall^\alpha i \in I)\; R^{\mathcal{M}_i}(\vec{a}(i))$$

We use the notation $\prod_{i\in I} \mathcal{M}_i/\alpha$ to denote the ultraproduct $\mathcal{L}$-structure $\mathcal{M}_\infty$.

It obvious that the interpretation of $\mathcal{L}$ for $\mathcal{M}_\infty$ is well-defined, i.e. does not depend on the choice of the representative $\vec{a}$ of its $=_\alpha$-equivalence class.

The following theorem shows that clause (iv) propagates over all $\mathcal{L}$-formulas in general, supporting the earlier made remark that the ultraproduct structure $\mathcal{M}_\infty$ can, indeed, be viewed as a limit of the sequence of structures $(M_i)_{i\in I}$ along the ultrafilter α.

Theorem 2.3 (Łoś). *Let $\mathcal{M}_\infty$ be the ultraproduct $\mathcal{L}$-structure of $\mathcal{L}$-structures $(M_i)_{i\in I}$ along an ultrafilter α. For each $\mathcal{L}$-formula $\varphi(\vec{x})$ and for each vector $\vec{a} \in (\prod_{i\in I} M_i)^{|\vec{x}|}$,*

$$\mathcal{M}_\infty \vDash \varphi([\vec{a}]_\alpha) \iff (\forall^\alpha i \in I)\, \mathcal{M}_i \vDash \varphi(\vec{a}(i)).$$

Proof. We prove by induction on the complexity of the formula $\varphi(\vec{x})$. The base case for relations is by definition. The case of the connective $\neg$ follows from the De Morgan Law for ultrafilters mentioned above: $\neg\forall^\alpha i$ is equivalent $\forall^\alpha i \neg$. The closedness of the ultrafilter α under finite intersection immediately handles the case of the connective $\wedge$, so the only connective left to handle is $\exists$.

To this end, let $\varphi(\vec{x}) := \exists y \psi(\vec{x}, y)$. Suppose that $\mathcal{M}_\infty \vDash \varphi([\vec{a}]_\alpha)$, so there is $[b]_\alpha \in \mathcal{M}_\infty$ such that $\mathcal{M}_\infty \vDash \psi(\vec{a}, b)$ and applying the induction hypothesis finishes the left-to-right implication. For the other implication, suppose the right handside and let $J \subseteq I$ be the α-large set of all $i \in I$ for which $\mathcal{M}_i \vDash \exists y \psi(\vec{a}(i), y)$. Using the Axiom of Choice, for each $i \in J$, choose $b_i \in M_i$ with $\mathcal{M}_i \vDash \psi(\vec{a}(i), b_i)$, and for each $i \in I \setminus J$, choose any $b_i \in M_i$. Thus, we have

$$(\forall^\alpha i \in I)\, \mathcal{M}_i \vDash \psi(\vec{a}(i), b_i),$$

which finishes the proof. □

2.B. **The Compactness theorem**

Throughout this subsection, we fix a countable language $\mathcal{L}$. Everything below can be done for uncountable languages as well, but we stick with countable to avoid unnecessary set-theoretic complications.

Ultraproducts and Łoś's theorem give an immediate proof of what is often referred to as "the most useful theorem of logic".

Compactness Theorem 2.4 (Gödel, Maltsev). *For any first-order language $\mathcal{L}$, if an $\mathcal{L}$-theory T is finitely satisfiable, then it is satisfiable. In other words, if every finite subset of T has a model, then so does T.*

Proof. Even though the theorem is true for languages of arbitrary cardinality, we will only prove it for countable languages to avoid the involvement of set theory.

Thus, we assume that $\mathcal{L}$ is countable and hence so is T. Take an enumeration $T = (\varphi_n)_{n\in\mathbb{N}}$. Let $\mathcal{M}_i$ be a model of $T_i := \{\varphi_n\}_{n\leq i}$ and take an ultraproduct $\mathcal{M}_\infty$ of $(\mathcal{M}_i)_{i\in\mathbb{N}}$ along a nonprincipal ultrafilter α on $\mathbb{N}$. We show that $\mathcal{M}_\infty \vDash T$. Fix $n \in \mathbb{N}$ in order to show that $\mathcal{M}_\infty \vDash \varphi_n$. By Łoś's theorem, all we need to show is that

$$(\forall^\alpha i \in \mathbb{N})\, \mathcal{M}_i \vDash \varphi_n.$$

By the choice of the $\mathcal{M}_i$,

$$(\forall i \geq n)\, \mathcal{M}_i \vDash \varphi_n,$$

and because α is nonprincipal, the set $\{i \in \mathbb{N} : i \geq n\}$ is α-large, so we are done. □

The following statement is equivalent to the Compactness theorem (via the contrapositive), but it gives a different way of looking at it thus enriching the prospect of applications.

Compactness Theorem 2.5 (Finite base version). *For a theory T and sentence φ in the language $\mathcal{L}$, if $T \vDash \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \vDash \varphi$.*

We leave the proof of the equivalence of the last two theorems as an exercise.

Topological version. The term compactness suggests that perhaps the Compactness theorem is equivalent to a statement that some topological space is compact. This is indeed the case and we proceed with the description of this space.

For a fixed language $\mathcal{L}$, let $\mathfrak{T}$ denote the set of all satisfiable *fully complete*[4] $\mathcal{L}$-theories and we equip this set with the topology generated by the sets $\langle\varphi\rangle := \{T \in \mathfrak{T} : T \vDash \varphi\}$.

It is easy to see that the sets $\langle\varphi\rangle$ form an algebra. In particular, they form a basis for the topology, making it zero-dimensional[5] Hausdorff.

Compactness Theorem 2.6 (Topological version)**.** *The topological space $\mathfrak{T}$ is compact.*

It shouldn't be too hard for readers familiar with pointset topology to prove that this version of the Compactness theorem is equivalent to the finite base version stated above. Instead of using the "open covers" definition of compactness, one should use its equivalent dual version involving closed sets, namely:

Proposition 2.7. *A topological space is compact if and only if every family of closed sets with the finite intersection property*[6] *has a nonempty intersection.*

The Compactness theorem, just like any compactness statement in general, provides a two-way bridge between the finite and the infinite. We proceed with applications illustrating this phenomenon.

2.C. **From finite to infinite**

In the previous course, you have discussed how the Compactness theorem yields statements like

(a) If a theory has arbitrarily large finite models then it also has an infinite model.
(b) If every finite subgraph of a graph admits a k-coloring, $k \in \mathbb{N}$, then so does the whole graph.

Here are a couple more examples of the same kind.

Let G be an (undirected) locally finite graph. A coloring $c : V(G) \to \{0,1\}$ is called *unfriendly* if for each vertex $v \in V(G)$ at least half of its neighbors have a different color from v. Think of this coloring as a partition of $V(G)$ into two political parties such that the majority of neighbors of each person belong to the opposite party.

Theorem 2.8. *Every locally finite graph admits an unfriendly coloring.*

This would immediately follow by the Compactness theorem once the following is proven, and we will the details as an exercise:

Lemma 2.9. *Every finite graph admits an unfriendly coloring.*

Proof. Left as an exercise. □

Another, slightly more involved application is the following.

Theorem 2.10 (Łoś, Marczewski)**.** *Let $\mathcal{A} \subseteq \mathcal{B}$ be algebras on a set X. Any finitely additive measure $\mu_{\mathcal{A}}$ admits an extension $\mu_{\mathcal{B}}$ to a finitely additive measure on $\mathcal{B}$.*

[4]An $\mathcal{L}$-theory T is called *fully complete* if for every $\mathcal{L}$-sentence φ, either $\varphi \in T$ or $(\neg\varphi) \in T$.

[5]A topology is called *zero-dimensional* if it has a basis consisting of clopen (i.e. both closed and open) sets.

[6]A family $\mathcal{F}$ of sets is said to have the *finite intersection property* if for every finite $\mathcal{F}_0 \subseteq \mathcal{F}$, $\bigcap_{A \in \mathcal{F}_0} A \neq \varnothing$.

This theorem will follow by an application of the Compactness theorem and its finite version:

Lemma 2.11. *Let $\mathcal{A} \subseteq \mathcal{B}$ be finite algebras on a set X. Any finitely additive measure $\mu_{\mathcal{A}}$ admits an extension $\mu_{\mathcal{B}}$ to a finitely additive measure on $\mathcal{B}$.*

Proof. Outlined in exercises. □

To apply the Compactness theorem to the last lemma, we need to figure out what first-order language to use. It is the underlying sets of our structures should subsets of $\mathcal{B}$ (unfortunate formatting conflict as our convention is that calligraphic letters denote the structure and not the underlying set), but the difficulty is that measures are real-valued functions, whereas functions and relations a subset of $\mathcal{B}$ aren't. The idea is to imitate real-valued functions by a bunch of unary relations!

Proof of Theorem 2.10. We define a language $\mathcal{L}$ as follows: for each $B \in \mathcal{B}$, put a constant symbol c_B in $\mathcal{L}$; furthermore, for each non-negative real s (taking only rationals would be enough too), put a unary predicate R_s in $\mathcal{L}$. What we have in mind is interpreting (informally)

$$R_s(c_B) :\Leftrightarrow \text{the measure of } B \text{ is at least } s.$$

This intuition leads us to defining an $\mathcal{L}$-theory T as follows: for $A, B \in \mathcal{B}$ and $s, t \in [0, +\infty)$,

(i) '$R_0(c_B)$' $\in T$;
(ii) if $A \in \mathcal{A}$ and $s \leq \mu_{\mathcal{A}}(A) < t$, then '$R_s(c_A)$' $\in T$ and '$\neg R_t(c_A)$' $\in T$;
(iii) if $s \leq t$ then '$R_t(c_B) \to R_s(c_B)$' $\in T$;
(iv) if $A \cap B = \varnothing$ then '$\big(R_s(c_A) \wedge R_t(c_B)\big) \to R_{s+t}(c_{A\cup B})$' $\in T$ and '$\big(\neg R_s(c_A) \wedge \neg R_t(c_B)\big) \to \neg R_{s+t}(c_{A\cup B})$' $\in T$.

Clearly, any extension $\mu_{\mathcal{B}}$ of $\mu_{\mathcal{A}}$ to $\mathcal{B}$ yields a model of T. Conversely, given a model $\mathcal{M}$ of T, we define $\mu_{\mathcal{B}}$ on $\mathcal{B}$ as follows: for each $B \in \mathcal{B}$,

$$\mu_{\mathcal{B}}(B) := \sup \{s \in [0, +\infty) : \mathcal{M} \models R_s(c_B)\}.$$

This is well-defined due to (i). In fact, it shouldn't be hard to verify that $\mu_{\mathcal{B}}$ is an f.a. measure extending $\mu_{\mathcal{A}}$. Thus, it only remains to prove that T is satisfiable, which is immediate by the Compactness theorem and Lemma 2.11. □

2.D. **From infinite to finite**

In arithmetic combinatorics and Ramsey theory, it often happens that one proves an infinitary theorem (e.g. theorems of Ramsey, van der Waerden, Szemerédi, etc.) by infinitary means (i.e. idealistic tools, without keeping track of ε's and bounding errors) and then deduces its finitary version via a so-called *compactness-and-contradiction* argument. The latter uses the fact that product of finite topological spaces is compact by Tychonoff's theorem. Here we give an example of such a proof using the Compactness theorem rather than a compactness-and-contradiction argument. Our example will be the deduction of the finite Ramsey theorem from its famous infinite counterpart. An analogous deduction of the finite version of van der Waerden's theorem from the infinite version is outlined in the exercises.

Put $\mathbf{n} := \{0, 1, \cdots, n-1\}$.

Theorem 2.12 (Finite Ramsey). *For every (number of colors) $k \geq 2$ and (desired size of a monochromatic set) $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that for any k-coloring χ of $[\mathbf{n}]^2$, there exists a χ-monochromatic subset $A \subseteq \mathbf{n}$ of cardinality m.*

Proof. Let $\mathcal{L}$ be the language containing constant symbols c_n, for every $n \in \mathbb{N}$, and unary relation symbols R_i, for every $i < k$. We think of R_i as a symbol for the color i, i.e. the color of $\{x, y\}$ is i if $R_i(x, y)$. This is easily expressed in an $\mathcal{L}$-sentence ψ that states that for every x exactly one R_i holds.

Fix $m \in \mathbb{N}$, and for each $n \in \mathbb{N}$, let φ_n be an $\mathcal{L}$-sentence expressing that $c_0, c_1, ..., c_{n-1}$ are pairwise distinct and the set $\{c_0, c_1, ..., c_{n-1}\}$ does not have a monochromatic subset of cardinality m (there are only finitely many such subsets, so we can express it).

Now suppose towards a contradiction that for any n, there is a k-coloring of $[\mathbf{n}]^2$ such that $\mathbf{n}$ has no monochromatic subsets of cardinality m. Thus, the theory $T := \{\psi\} \cup \{\varphi_n : n \in \mathbb{N}\}$ is finitely satisfiable, and hence, has a model $\mathcal{M}$. Let $C := \{c_n^{\mathcal{M}} : n \in \mathbb{N}\}$. By the Infinite Ramsey theorem, C has an infinite monochromatic subset A, i.e. there is $i < k$ such that for all distinct $a, a' \in A$, $R_i^{\mathcal{M}}(a, a')$. Let n be large enough so that $A \cap \{c_i : i < n\}$ has at least m elements. Then it is clear that $\mathcal{M} \not\models \varphi_n$, a contradiction. $\square$

The original combinatorial proof of this is much messier (look it up).

2.E. **Saturation and ultraproducts**

In this subsection, we return to ultraproducts and prove their most important property, namely, *countable saturation*, which is what makes them so useful.

Throughout this subsection, fix a language $\mathcal{L}$. For an $\mathcal{L}$-structure $\mathcal{M}$ and a parameter set $A \subseteq M$, let $\mathfrak{D}^n_{\mathcal{M}}(A)$ denote the collection of A-definable subsets of M^n and put $\mathfrak{D}_{\mathcal{M}}(A) := \bigsqcup_{n \in \mathbb{N}} \mathfrak{D}^n_{\mathcal{M}}(A)$.

Proposition 2.13. *For each $n \geq 1$, $\mathfrak{D}^n_{\mathcal{M}}(A)$ is an algebra. Moreover, $\mathfrak{D}_{\mathcal{M}}(A)$ is closed under projections.*

Proof. Complements, unions, and projections correspond to $\neg$, $\wedge$, and $\exists$, respectively. $\square$

Definition 2.14. For a cardinal κ, an $\mathcal{L}$-structure $\mathcal{M}$ is called *κ-saturated* if for every $n \geq 1$ and $A \subseteq M^n$ with $|A| < \kappa$, any family of sets from $\mathfrak{D}^n_{\mathcal{M}}(A)$ with the finite intersection property has nonempty intersection. When $\kappa = \aleph_1$, i.e. A ranges over countable sets, we often use the term *countably saturated* instead of $\aleph_1$-saturated.

For the readers familiar with pointset topology, we give a topological reformulation:

Proposition 2.15. *For a cardinal κ, an $\mathcal{L}$-structure $\mathcal{M}$ is κ-saturated if and only if for every $n \geq 1$ and $A \subseteq M^n$ with $|A| < \kappa$, the topology on M^n generated by $\mathfrak{D}^n_{\mathcal{M}}(A)$ is compact.*

Proof. Immediate using Proposition 2.7. $\square$

The following proposition shows that, somewhat surprisingly, it is enough to check saturation for one-dimensional sets.

Proposition 2.16. *For a cardinal κ, an $\mathcal{L}$-structure $\mathcal{M}$ is κ-saturated if and only if for every $A \subseteq M^1$ with $|A| < \kappa$, any family of sets from $\mathfrak{D}^1_{\mathcal{M}}(A)$ with the finite intersection property has nonempty intersection.*

Proof. We prove the nontrivial implication by induction on the dimension n. Suppose the statement is true for all $k < n$, $n \geq 2$. We'd like to reduce dimension so that the induction hypothesis kicks in. There are, in general, two ways to reduce dimension: projections and taking fibers. Here, one has to use both (in this order) and we will leave this as an exercise. $\square$

It is not even clear a priori that κ-saturated structures exist. However, such a structure can be built as a union of an increasing sequence of richer and richer elementary extensions, which is obtained by iterative applications of the Compactness theorem. Instead of working this out in detail, we will give a nice proof for $\kappa := \aleph_1$ using ultraproducts.

Theorem 2.17. *Let $\mathcal{L}$ be a countable language and α a nonprincipal ultrafilter on $\mathbb{N}$. The ultraproduct $\mathcal{M}_\infty$ over α of any sequence $(\mathcal{M}_i)_{i\in\mathbb{N}}$ of $\mathcal{L}$-structures is countably saturated.*

Proof. By Proposition 2.16, it is enough to check saturation for one-dimensional sets, so fix a countable parameter set $A \subseteq M_\infty$ and let $\mathfrak{B} \subseteq \mathfrak{D}^1_{\mathcal{M}_\infty}(A)$ have the finite intersection property. Because $\mathcal{L}$ and A are countable, $\mathfrak{B}$ is also countable and we take an enumeration $\mathfrak{B} = \{B^{(n)}\}_{n\in\mathbb{N}}$. We need to show that $\bigcap \mathfrak{B} := \bigcap_{n\in\mathbb{N}} B^{(n)}$ is nonempty.

By one of the exercises (or, basically, by Łoś's theorem), each $B^{(n)}$ is a quasibox, i.e.

$$B^{(n)} := \left[\prod_{i\in\mathbb{N}} B_i^{(n)}\right]_\alpha,$$

and it doesn't matter for the rest of the proof that the sets $B_i^{(n)}$ are definable, so we can forget about definability.

Claim. For each $N \in \mathbb{N}$, we have $(\forall^\alpha i \in \mathbb{N}) \bigcap_{n\leq N} B_i^{(n)} \neq \varnothing$.

Proof of Claim. Fixing $N \in \mathbb{N}$, the finite intersection property of $\mathcal{B}$ gives $\bigcap_{n\leq N} B^{(n)} \neq \varnothing$, so there is $x \in \bigcap_{n\leq N} B^{(n)}$. Thus, $(\forall n \leq N)\ x \in \left[\prod_{i\in\mathbb{N}} B_i^{(n)}\right]_\alpha$, which means that $(\forall n \leq N)(\forall^\alpha i \in \mathbb{N})\ x(i) \in B_i^{(n)}$. By the closedness of α under finite intersections, we may switch the quantifiers $(\forall n \leq N)$ and $(\forall^\alpha i \in \mathbb{N})$, obtaining $(\forall^\alpha i \in \mathbb{N})(\forall n \leq N)\ x(i) \in B_i^{(n)}$. This means that $(\forall^\alpha i \in \mathbb{N})\ x(i) \in \bigcap_{n\leq N} B_i^{(n)}$, so, in particular, for α-a.e. $i \in \mathbb{N}$, the set $\bigcap_{n\leq N} B_i^{(n)}$ is nonempty. $\dashv$

We are now ready to define an element x of $\bigcap_{n\in\mathbb{N}} B^{(n)}$. For each $i \in \mathbb{N}$, let N_i denote the largest natural number $\leq i$ such that $\bigcap_{n\leq N_i} B^{(n)} \neq \varnothing$. Using the Axiom of Choice, take $x_i \in \bigcap_{n\leq N_i} B^{(n)}$ and define $x := [(x_i)_{i\in\mathbb{N}}]_\alpha$.

Fixing an arbitrary $N \in \mathbb{N}$, it remains to show that $x \in B^{(N)}$, or equivalently,

$$(\forall^\alpha i \in \mathbb{N})\ x(i) \in B_i^{(N)}.$$

But by the claim above and the definition of N_i, $(\forall^\alpha i \in \mathbb{N})\ N_i \geq \min(N, i)$, and because α is nonprincipal, we also have $(\forall^\alpha i \in \mathbb{N})\ i \geq N$, or, in other words, $(\forall^\alpha i \in \mathbb{N})\ \min(N, i) = N$. Combining the two α-a.e. statements together (using the closedness of α under finite intersections) gives $(\forall^\alpha i \in \mathbb{N})\ N_i \geq N$, which implies that, in fact,

$$(\forall^\alpha i \in \mathbb{N})\ x(i) \in \bigcap_{n\leq N} B_i^{(n)}.$$

$\square$

2.F. **Ultrapowers as saturated extensions**

Let $\mathcal{L}$ be a fixed language, an ultrafilter α on a set I, and an $\mathcal{L}$-structure $\mathcal{M}$. The ultraproduct $\prod_{i \in I} \mathcal{M}_i$ of the constant sequence $(\mathcal{M})_{i \in I}$ is called the *ultrapower of* $\mathcal{M}$ *along* α and denoted by $\mathcal{M}^I/\alpha$.

Theorem 2.18. *The diagonal map* $\mathcal{M} \hookrightarrow \mathcal{M}^I/\alpha$ *by* $a \mapsto [(a)_{i \in I}]_\alpha$ *is an elementary* $\mathcal{L}$*-embedding.*

Proof. Easily follows from Łoś's theorem. □

The last theorem together with Theorem 2.17 yields.

Corollary 2.19. *For any countable language* $\mathcal{L}$*, every* $\mathcal{L}$*-structure* $\mathcal{M}$ *admits a countably saturated elementary extension, namely its ultrapower along a nonprincipal ultrafilter on* $\mathbb{N}$*.*

This corollary is actually true more generally:

Proposition 2.20. *For any language* $\mathcal{L}$ *and any cardinal* κ*, every* $\mathcal{L}$*-structure admits a* κ*-saturated elementary extension.*

Although we won't prove the last proposition in these notes, we will use it below to build nonstandard extensions for uncountable languages.

3. Nonstandard analysis

Now we are ready to build an elementary extension $\mathbb{R}^*$ of $\mathbb{R}$ that inherits enough of the structure and properties of $\mathbb{R}$ and yet is countably saturated, whence contains idealistic elements such as infinitesimals.

3.A. **Hyperreals**

Let $\mathcal{L}_{of}$ be the language of ordered fields, i.e. $(0, 1, +, \cdot, <)$. We extend this language to $\mathcal{L}$ by adding the following:

(i) a constant symbol c_r for every $r \in \mathbb{R}$;
(ii) an n-ary relation symbol P_A for every $A \subseteq \mathbb{R}^n$ and $n \geq 1$;
(iii) an n-ary function symbol F_f for every function $f : \mathbb{R}^n \to \mathbb{R}$ and $n \geq 1$.

Let $\mathcal{R}$ be the $\mathcal{L}$-structure with the underlying set $\mathbb{R}$ and natural (standard) interpretation of $\mathcal{L}$. By Proposition 2.20, $\mathcal{R}$ has a countably saturated elementary extension $\mathcal{R}^*$.

Notation 3.1. We denote the underlying set of $\mathcal{R}^*$ by $\mathbb{R}^*$. Moreover, for every $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$, put $A^* := P_A^{\mathcal{R}^*}$ and $f^* := F_f^{\mathcal{R}^*}$. Call A^* and f^* the *nonstandard extensions* of A and f, respectively.

By elementarity, $\mathcal{R}^*$ has the following properties:

(NS1) The reduct of $\mathcal{R}^*$ to $\mathcal{L}_{of}$ is an ordered field.
(NS2) For every $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$, let $f_1 : \mathbb{R}^n \to \mathbb{R}$ be any extension of f and put $f^* := f_1^* \downharpoonright_{A^*}$. The function $f^* : A^* \to \mathbb{R}^*$ is well-defined, i.e. is independent of the choice of the extension f_1 of f. We again call f^* the *nonstandard extension* of f.
(NS3) Transfer Principle: for every $\mathcal{L}$-sentence φ, $\mathcal{R} \models \varphi$ if and only if $\mathcal{R}^* \models \varphi$.

Moreover, countable saturation gives:

(NS4) Existence of infinitesimals: $\mathbb{R}^*$ has a positive *infinitesimal* element ε, i.e. $\varepsilon > 0$ and $\varepsilon < \frac{1}{n}$ for all $n \in \mathbb{N}$.

The elements of $\mathbb{R}^*$ are called *hyperreals* and we refer to $\mathbb{R}^*$ as the *ordered field of hyperreals*. Henceforth, we abandon the notation $\mathcal{R}$ and $\mathcal{R}^*$ and use $\mathbb{R}$ and $\mathbb{R}^*$ for both the structures and the underlying sets. We also call $\mathbb{R}^*$ a *nonstandard extension of* $\mathbb{R}$.

3.B. **Arithmetic in $\mathbb{R}^*$**

Let ε be a positive infinitesimal.

- $-\varepsilon$ is a negative infinitesimal.
- $r\varepsilon$ is an infinitesimal for every $r \in \mathbb{R}$.
- ε^{-1} is a positive *infinite* element, i.e. $\varepsilon^{-1} > n$ for every $n \in \mathbb{N}$. Consequently, $-\varepsilon^{-1}$ is a negative infinite element.

We make all these terms more precise.

Definition 3.2.

(a) The set of *finite hyperreals* is $\mathbb{R}_{\text{fin}} := \{x \in \mathbb{R}^* : |x| \leq n \text{ for some } n \in \mathbb{N}\}$.

(b) The set of *infinite hyperreals* is $\mathbb{R}_{\text{inf}} := \mathbb{R}^* \setminus \mathbb{R}_{\text{fin}}$.

(c) The set of *infinitesimal hyperreals* is $\mu := \left\{x \in \mathbb{R}^* : |x| < \frac{1}{n} \text{ for all } n \in \mathbb{N}\right\}$.

Proposition 3.3.

(a) $\mathbb{R}_{\text{fin}}$ *is a subring of* $\mathbb{R}^*$.

(b) μ *is an ideal in* $\mathbb{R}_{\text{fin}}$.

Proof. Left as an exercise. □

A natural question now arises: What is the quotient ring $\mathbb{R}_{\text{fin}}/\mu$? The answer will arrive shortly.

Definition 3.4. For $x, y \in \mathbb{R}^*$, say that x and y are *infinitely close*, and write $x \approx y$, if $x - y \in \mu$.

Clearly, $\approx$ is an equivalence relation, and in fact, it is a *congruence relation*, i.e. $x \approx y$ and $u \approx v$ implies $x \pm u \approx y \pm v$.

Proposition 3.5 (Existence of standard parts)**.** *For every $r \in \mathbb{R}_{\text{fin}}$, there is a unique $s \in \mathbb{R}$ such that $r \approx s$. We call s the standard part of r and write* $\text{st}(r) = s$.

Proof. The uniqueness is obvious and we show existence. Without loss of generality, we can assume $r > 0$. Because $r \in \mathbb{R}_{\text{fin}}$, the set

$$A := \{a \in \mathbb{R} : a < r\}$$

is bounded above as a subset of $\mathbb{R}$, so by the completeness of $\mathbb{R}$, $s := \sup(A)$ exists and it is easy to see that $s \approx r$. □

Proposition 3.6. *The map* $\text{st} : \mathbb{R}_{\text{fin}} \to \mathbb{R}$ *is a ring homomorphism.*

Proof. Left as an exercise. □

Corollary 3.7. $\mathbb{R}_{\text{fin}}/\mu \cong \mathbb{R}$. *In particular, μ is a maximal ideal of* $\mathbb{R}_{\text{fin}}$.

Proof. The kernel of st is μ, so $\mathbb{R}_{\text{fin}}/\mu \cong \mathbb{R}$ by the First Isomorphism theorem. Because $\mathbb{R}$ is a field, μ is a maximal ideal. □

3.C. **Order structure of $\mathbb{R}^*$**

Proposition 3.8. *$\mathbb{N}^*$ is cofinal in $\mathbb{R}^*$, i.e. for every $x \in \mathbb{R}^*$ there is $N \in \mathbb{N}^*$ such that $N \geq x$. In particular, $\mathbb{N}^* \setminus \mathbb{N} \neq \varnothing$.*

Proof. The first statement follows immediately by Transfer. Thus, because $\mathbb{R}^*$ has positive infinite elements, $\mathbb{N}^*$ must also have infinite elements, and hence, $\mathbb{N}^* \setminus \mathbb{N} \neq \varnothing$. □

Notation 3.9. We write $N > \mathbb{N}$ to mean $N \in \mathbb{N}^* \setminus \mathbb{N}$.

Because $\mathbb{R}_{\text{fin}}$ is a subgroup of the (abelian) group $\mathbb{R}^*$ under addition, we can let $\sim_{\text{fin}}$ denote the coset equivalence relation. For each $x \in \mathbb{R}^*$, we denote by $[x]_{\text{fin}}$ the coset of x and call it the *Archimedean class* of x. The Archimedean class $[0]_{\text{fin}} = \mathbb{R}_{\text{fin}}$ is called *finite*; the other Archimedean classes are called *infinite*.

Note that the relation $\sim_{\text{fin}}$ respects $<$, i.e. if $x \sim_{\text{fin}} x' \nsim_{\text{fin}} y \sim_{\text{fin}} y'$, then $x < y$ if and only if $x' < y'$. This allows us to define a linear ordering on the Archimedean classes:

$$[x]_{\text{fin}} \leq [y]_{\text{fin}} :\Leftrightarrow x \leq y.$$

As usual, we write $[x]_{\text{fin}} < [y]_{\text{fin}}$ if $[x]_{\text{fin}} \leq [y]_{\text{fin}}$ and $[x]_{\text{fin}} \neq [y]_{\text{fin}}$. Call an Archimedean class $[x]_{\text{fin}}$ positive if $[x]_{\text{fin}} > [0]_{\text{fin}}$, and negative if $[x]_{\text{fin}} < [0]_{\text{fin}}$.

Perhaps somewhat surprisingly, this ordering on the Archimedean classes is not at all discrete as the following proposition shows.

Proposition 3.10. *The ordering $<$ on the positive (resp. negative) infinite Archimedean classes is a dense linear ordering without endpoints.*

Proof. Left as an exercise. □

3.D. **Nondefinable subsets of $\mathbb{R}^*$**

Proposition 3.11. *The sets $\mathbb{N}$ and $\mathbb{R}$ are not definable in $\mathbb{R}^*$.*

Proof. Because $\mathbb{N} = \mathbb{N}^* \cap \mathbb{R}$, the nondefinability of $\mathbb{N}$ implies that of $\mathbb{R}$, so we only need to show that the former. Suppose for contradiction that $\mathbb{N}$ is definable in $\mathbb{R}^*$ by $\varphi(x, \vec{a})$, where $\varphi(x, \vec{y})$ is an $\mathcal{L}$ formula and $\vec{a} \in (\mathbb{R}^*)^{|\vec{x}|}$. Then we have $\mathbb{R}^* \models \varphi(0, \vec{a})$ and $\mathbb{R}^* \models \forall(n \in \mathbb{N}^*)\varphi(n, \vec{a}) \to \varphi(n+1, \vec{a})$. Because induction holds in $\mathbb{N}$, it also holds in $\mathbb{N}^*$ by Transfer, yielding $\mathbb{R}^* \models \forall(n \in \mathbb{N}^*)\varphi(n, \vec{a})$ and hence $\mathbb{N}^* = \mathbb{N}$, contradicting Proposition 3.8. □

3.E. **Nonstandard calculus**

Sequences. Viewing a sequence $(s_n)_{n \in \mathbb{N}}$ as a function $s : \mathbb{N} \to \mathbb{R}$, it makes sense to talk about its nonstandard extension $s^* : \mathbb{N}^* \to \mathbb{R}^*$ and, abusing notation, we write $s_N := s^*(N)$ for $N > \mathbb{N}$.

Proposition 3.12. *For a (standard) sequence $(s_n)_{n \in \mathbb{N}}$ and $L \in \mathbb{R}$, $(s_n)_{n \in \mathbb{N}} \to L$ if and only if $s_N \approx L$ for all $N > \mathbb{N}$.*

Proof. $\Rightarrow$: Fix an arbitrary real $\varepsilon > 0$. Then there is $m \in \mathbb{N}$ such that for each $n \geq m$, $|s_n - L| < \varepsilon$. Transferring this with m fixed, we get that in $\mathbb{R}^*$, the following holds: for each $N > m$, $|s_N - L| < \varepsilon$. But this implies that for each $N > \mathbb{N}$, $|s_N - L| < \varepsilon$. Because ε is arbitrary, we get $s_N \approx L$.

$\Leftarrow$: Fix a real $\varepsilon > 0$. It is true in $\mathbb{R}^*$ that there is $N \in \mathbb{N}^*$ (namely, any infinite N) such that for every $n > N$, $s_n \approx L$; in particular, $|s_n - L| < \varepsilon$. Transfering this back to $\mathbb{R}$ gives: $\exists N \in \mathbb{N} \forall n > N \; |s_n - L| < \varepsilon$. ☐

Continuity. Henceforth, when considering the nonstandard extension f^* of a function f, we drop the * and simply write f (just like we did with sequences).

Proposition 3.13. *For $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and $a \in A$, the following are equivalent:*
(1) *f is continuous at a;*
(2) *if $x \in A^*$ and $x \approx a$, then $f(x) \approx f(a)$;*
(3) *there is a positive $\delta \in \mu$ such that, for all $x \in A^*$, if $|x - a| < \delta$, then $f(x) \approx f(a)$.*

Proof. (1)$\Rightarrow$(2): Fixing a real $\varepsilon > 0$, we need to show that whenever $A^* \ni x \approx a$, we have $|f(x) - f(a)| < \varepsilon$. But by (1), there is a real $\delta > 0$ such that

$$\mathbb{R} \models \forall x \in A(|x - a| < \delta \to |f(x) - f(a)| < \varepsilon),$$

and transferring this gives

$$\mathbb{R}^* \models \forall x \in A^*(|x - a| < \delta \to |f(x) - f(a)| < \varepsilon).$$

If $x \approx a$, then in particular $|x - a| < \delta$, which gives $|f(x) - f(a)| < \varepsilon$.

(2)$\Rightarrow$(3): Trivial.

(3)$\Rightarrow$(1): For an arbitrary real $\varepsilon > 0$, condition (3) in particular gives

$$\mathbb{R}^* \models \exists \delta > 0 \forall x \in A^*(|x - a| < \delta \to |f(x) - f(a)| < \varepsilon),$$

transferring which gives (1). ☐

The following shows the subtle difference between continuity and uniform continuity.

Proposition 3.14. *$f : A \to \mathbb{R}$ is uniformly continuous if and only if for all $x, y \in A^*$, if $x \approx y$ then $f(x) \approx f(y)$.*

Proof. Left as an exercise. ☐

Remark 3.15. Thus, the difference between continuity and uniform continuity is that in the former case, one of the points is always standard, while in the latter both x, y can be nonstandard.

Intermediate Value Theorem 3.16. *Let $f : [a, b] \to \mathbb{R}$ be continuous. For every real d strictly in between $f(a)$ and $f(b)$, there is $c \in (a, b)$ with $f(c) = d$.*

Proof. We will use the so-called *hyperfinite method*: we will take an infinite $N > \mathbb{N}$ and partition $[a, b]$ into N-many subintervals, each of length $\frac{1}{N}$.

Without loss of generality, suppose $f(a) < d < f(b)$. Define a sequence (s_n) as follows: for $n > 0$, let $\{p_0; p_1, \cdots, p_n\}$ denote the partition of $[a, b]$ into n equal pieces of width $\frac{b-a}{n}$, so $p_0 = a$ and $p_n = b$. Since $f(p_0) < d$, there must be $s_n := \max\{p_k : f(p_k) < d\}$, so p_k is the "last time" that $f(p_k) < d$. Observe that $s_n < b$.

We now fix $N > \mathbb{N}$ and claim that $c := \mathrm{st}(s_N) \in [a, b]$ is as desired, namely, that $f(c) = d$. (Note that $s_N \in [a, b]$, whence $\mathrm{st}(s_N)$ is defined.) Indeed, by transfer, $s_N < b$, whence $s_N + \frac{b-a}{N} \leq b$. Again, by transfer,

$$f(s_N) < d < f(s_N + \frac{b-a}{N}).$$

However, $s_N + \frac{b-a}{N} \approx s_N \approx c$, so the continuity of f gives

$$f(c) \approx f(s_N) < d < f(s_N + \frac{b-a}{N}) \approx f(c),$$

whence $f(c) \approx d$. Since both $f(c)$ and d are reals, they must be equal. □

Limits and differentiation. In this subsection, we let $f : A \to \mathbb{R}$ and $a \in A$ be an interior point.

Proposition 3.17. *For $a \in A$ and $f : A \to \mathbb{R}$, $\lim_{x \to a} f(x) = L$ if and only if for all $x \in A^*$, if $x \approx a$ but $x \neq a$, then $f(x) \approx L$.*

Proof. Left as an exercise. □

Proposition 3.18. *f is differentiable at a with $f'(a) = D$ if and only if for every positive $\varepsilon \in \mu$, we have $\frac{f(a+\varepsilon)-f(a)}{\varepsilon} \approx D$.*

Proof. Immediate from Proposition 3.17. □

Suppose f is differentiable at a and fix a positive $dx \in \mu$ (we use the notation dx for nostalgic reasons). Putting $df := f(a + dx) - f(a)$, we see that $f'(a) \approx \frac{df}{dx}$, even through we would be scolded in a calculus class for treating $\frac{df}{dx}$ as an actual fraction!

Product Rule 3.19. *Suppose functions $f, g : A \to \mathbb{R}$ are differentiable at $x \in A$. Then $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.*

Proof. Fix positive $dx \in \mu$. Then

$$\begin{aligned} d(fg) &= f(x + dx)g(x + dx) - f(x)g(x) \\ &= (f(x) + df)(g(x) + dg) - f(x)g(x) \\ &= df \cdot g(x) + f(x) \cdot dg + df \cdot dg. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{d(fg)}{dx} &= \frac{df}{dx} g(x) + f(x)\frac{dg}{dx} + df \frac{dg}{dx} \\ &\approx f'(x)g(x) + f(x)g'(x) + df g'(x). \end{aligned}$$

By the continuity of f, $df \approx 0$, so we are done. □

Exercises

1. Show that the collection of nowhere dense subsets of a nonempty topological space X is an ideal.

2. Exhibit intermediate sets for ideals/filters given in Examples 1.2 and Example 1.6.

3. Show that the subsets of $\mathbb{N}$ for which the density is defined and is equal to 1 form a filter.

4. Fix a set X and prove that

(a) every ideal on X contains $\varnothing$ and every filter on X contains X;

(b) every algebra (resp. σ-algebra) on X contains both $\varnothing$ and X, and is closed under finite (resp. countable) intersections.

5. Let $\mathcal{B}$ be an algebra on a set X. Call a set $A \in \mathcal{B}$ an *atom* of $\mathcal{B}$ if it cannot be partition into two nonempty sets from $\mathcal{B}$, i.e. whenever A is a disjoint union of $B, C \in \mathcal{B}$, one of B, C is empty. Prove that if $\mathcal{B}$ is finite, then the set of its atoms is a partition of X.

6. Let X be a set, $\mathcal{A}$ an algebra on X, and μ a finitely additive measure on $\mathcal{A}$.

(a) Explicitly describe the algebra $\overline{\mathcal{A}}$ generated by $\mathcal{A}$ and all of the μ-null sets. More precisely, describe the sets in $\overline{\mathcal{A}}$ in terms of sets in $\mathcal{A}$ and μ-null sets.

(b) Show that μ can be uniquely extended to a finitely additive measure $\overline{\mu}$ on $\overline{\mathcal{A}}$.

(c) Show that $\overline{\mu}$ is complete.

7. Show that disjoint finite unions of intervals form an algebra on $\mathbb{R}$.

8. Prove that if $\mathcal{A} \subseteq \mathcal{B}$ are finite algebras on X, then any finitely additive measure $\mu_{\mathcal{A}}$ on $\mathcal{A}$ can be extended (typically not uniquely) to a finitely additive measure $\mu_{\mathcal{B}}$ on $\mathcal{B}$.

Hint: It is enough to define the value of $\mu_{\mathcal{B}}$ on the atoms of $\mathcal{B}$ and note that each $\mathcal{A}$-atom is a disjoint union of $\mathcal{B}$-atoms.

9. Work out the proof of König's Lemma 1.18 in detail.

10. Using the Infinite Ramsey Theorem 1.19, prove that every sequence of reals admits a monotone subsequence.

11. Let X_∞ be an ultraproduct of the sequence of sets $(A_i)_{i \in I}$ over some nonprincipal ultrafilter α. Call a set $A \subseteq X_\infty$ a *quasibox* (also known as *internal set*) if it is of the form $[\prod_{i \in I} A_i]_\alpha$, where $A_i \subseteq X_i$. Prove that quasi-boxes form an algebra.

Remark: This is perhaps somewhat counter-intuitive in comparison with the usual boxes (think of rectangles in $\mathbb{R}^2$—they do not form an algebra).

12. For an $\mathcal{L}$-structure $\mathcal{M}$ and an $\mathcal{L}$-formula $\varphi(\vec{x})$, let $\langle \varphi(\vec{x}) \rangle_{\mathcal{M}}$ denote the subset of $M^{|\vec{x}|}$ defined by $\varphi(\vec{x})$.

Let $\mathcal{M}_\infty$ be the ultraproduct $\mathcal{L}$-structure of the sequence $(\mathcal{M}_i)_{i \in I}$ of $\mathcal{L}$-structures. Show that for any $\mathcal{L}$-formula $\varphi(\vec{x})$, the set $\langle \varphi(\vec{x}) \rangle_{\mathcal{M}_\infty}$ is exactly the quasibox

$$\left[\prod_{i \in I} \langle \varphi(\vec{x}) \rangle_{\mathcal{M}_i} \right]_\alpha .$$

13. Prove the equivalence of all three of the forms of the Compactness theorem, namely: 2.4, 2.5, and 2.6.

14. This exercise isn't related to anything from lecture, but it is a useful tool to have and it will be used below. Prove the following statement:

∃-Elimination Rule. *Let c be a constant symbol not in a language $\mathcal{L}$. Let $\varphi(x)$ and ψ be an $\mathcal{L}$-formula and an $\mathcal{L}$-sentence, respectively. If $\varphi(c) \vDash \psi$ then $\exists x \varphi(x) \vDash \psi$.*

15. Verify that $\mu_{\mathcal{B}}$ defined in the proof of Theorem 2.6 is indeed a finitely additive measure extending $\mu_{\mathcal{A}}$.

16. Prove that any finite graph admits an unfriendly coloring and deduce the same for all locally finite graphs.

HINT: For a finite graph G, take a partition $V(G) = A \sqcup B$ (i.e. a 2-coloring) with $|(A,B)|$ being maximum possible, where (A,B) is the set of all edges between A and B (i.e. incident to both A and B). This partition is an unfriendly coloring.

17. The following is a well known theorem of additive combinatorics:

Theorem (van der Waerden)**.** *For any partition of $\mathbb{N}$ into finitely-many sets, one of these sets contains arbitrarily long arithmetic progressions.*

Use this theorem (without proof) and the Compactness theorem to derive the following finitary version:

Theorem (van der Waerden: finitary version)**.** *For any (number of sets in a partition) $k \geq 1$ and (desired length of arithmetic progressions) $l \geq 1$, there exists $n \in \mathbb{N}$ such that whenever $\mathbf{n} := \{0, 1, \cdots, n-1\}$ is partitioned into k sets, one of these sets contains an arithmetic progression of length l.*

18. Follow the steps below to show that the class $\mathfrak{D}$ of disconnected graphs is not axiomatizable. Assume for contradiction that there is an axiomatization T of $\mathfrak{D}$ in some language $\mathcal{L}$ containing a binary relation symbol E (edge-relation). Let $\mathcal{L}' := \mathcal{L} \cup \{u, v\}$, where u, v are new constant symbols and put

$$S := \{\chi_n(u,v) : n \in \mathbb{N}\},$$

where the formula $\chi_n(x,y)$ says that there is no path of length $\leq n$ between x and y (here x, y are variables) and also includes the axiom of being an undirected graph.

(i) Show that for every $\varphi \in T$ there is $n \in \mathbb{N}$ such that $\chi_n(u,v) \vDash \varphi$.

(ii) Conclude, using Exercise 14, that $\exists x \exists y \chi_n(x,y) \vDash \varphi$.

(iii) Put $S' := \{\exists x \exists y \chi_n(x,y) : n \in \mathbb{N}\}$ and conclude that $S' \vDash T$, i.e. for every $\varphi \in T$, $S' \vDash \varphi$. Explain why this is a contradiction.

19. Recall the language of graphs: $\mathcal{L}_{\text{graph}} := (E)$, where E is a binary relation symbol. Show that the relation

$$P(x,y) \iff x \text{ and } y \text{ are connected}$$

is not 0-definable in the undirected graph $\mathcal{G} := (G; E)$ that consists of two bi-infinite paths; more precisely $\mathcal{G}$ is a 2-regular[7] acyclic graph with two connected components.

[7] A graph is called *k-regular* if every vertex has exactly k neighbors.

Solution 1: Let $A, B \subseteq G$ denote the two connected components. Suppose for contradiction that there is an $\mathcal{L}_{\text{graph}}$-formula $\varphi(x, y)$ defining the relation P in $\mathcal{G}$. Using the Compactness theorem, get an elementary extension of $\mathcal{G}$ containing at least one (possibly more) other connected component C (necessarily a bi-infinite path) such that φ holds between the elements of A and C. But swapping B and C is an automorphism of this extended graph, so φ must also hold between the elements of A and B, contradicting the fact that the extension is elementary.

Solution 2: Prove that the theory of 2-regular acyclic graphs is uncountably categorical and hence complete. Therefore, there is no first-order difference between the graphs with one bi-infinite path and with two bi-infinite paths.

20. Work out the proof of Proposition 2.16 using the given outline.

21. Prove Proposition 3.3.

22. Show that $\approx$ is an equivalence relation on $\mathbb{R}^*$, and in fact, it is a *congruence relation*, i.e. $x \approx y$ and $u \approx v$ implies $x \pm u \approx y \pm v$.

23. Show that $\mathbb{R}^*$ is not a complete linear order, i.e. it has a bounded subset, namely $\mathbb{R}$, for which sup does not exist.

24. Prove Proposition 3.10.

25. Prove that a sequence $(s_n)_{n\in\mathbb{N}}$ is bounded in $\mathbb{R}$ if and only if $s_N \in \mathbb{R}_{\text{fin}}$ for all $N \in \mathbb{N}^*$.

26. Prove that $f : A \to \mathbb{R}$ is uniformly continuous if and only if for all $x, y \in A^*$, if $x \approx y$ then $f(x) \approx f(y)$.

27. For $a \in A$ and $f : A \to \mathbb{R}$, prove that $\lim_{x\to a} f(x) = L$ if and only if for all $x \in A^*$, if $x \approx a$ but $x \neq a$, then $f(x) \approx L$.

28. Prove the Chain Rule using nonstandard analysis.

References

[Gol] I. Goldbring, *Lecture notes on nonstandard analysis*, UCLA Summer School in Logic.

[Tse13] A. Tserunyan, *First order logic and Gödel incompleteness* (2013), unpublished lecture notes, available at `http://www.math.uiuc.edu/~anush/Notes/logic_lectures.pdf`.

Department of Mathematics, University of Illinois at Urbana-Champaign, IL, 61801, USA
E-mail address: anush@illinois.edu